PRIFYSGOL ABERYSTWYTH UNIVERSITY

www.aber.ac.uk

APPENDIX G

Statistical modelling of faecal indicator organisms at a marine bathing water site: results of an intensive study at Swansea Bay, UK

smart coasts

Volume 1 - Report

Smart Coasts - Sustainable Communities

ERDF
Llywodraeth Cynulliad Cymru Welsh Assembly Government
Ewrop & Chymru: Buddsoddi yn eich Dyfodol
Cronfa Datblygu Rhanbarthol Ewrop
Europe & Wales: Investing in your Future
European Regional Development Fund

SOUTHERN & EASTERN Regional Assembly
Promoting Our Region

IRELAND WALES 2007 – 2013

Ireland's EU Structural Funds Programmes 2007 - 2013
Co-funded by the Irish Government and the European Union

EUROPEAN REGIONAL DEVELOPMENT FUND

Statistical modelling of faecal indicator organisms at a marine bathing water site: results of an intensive study at Swansea Bay, UK

A report from the Interreg 4a Smart Coasts – Sustainable Communities Project

Mark D. Wyer[1], David Kay[1], Huw Morgan[2], Sam Naylor[2], Phil Govier[2], Simon Clark[2], John Watkins[1], Cheryl M. Davies[1], Carol Francis[1], Hamish Osborn[3] and Sarah Bennett[3]

August 2013

1 Institute of Geography and Earth Sciences
Llandinam Building
Aberystwyth University
SY23 3DB
UK

2 Environment Department
City and County of Swansea
The Guildhall
Swansea
SA1 4PE
UK

3 Natural Resources Wales
Area Office
Maes Newydd
Llandarcy
SA10 6JQ
UK

Summary

An investigation the feasibility of statistical modelling of faecal indicator organism concentrations (FIOs – *Escherichia coli* and intestinal enterococci (IE) as colony forming units (cfu)/100 ml)) at Swansea Bay, south Wales, UK is reported. Predictive modelling can mitigate bathing prohibition at sites that are unlikely to comply with the revised European bathing waters Directive (rBWD) by 2020. Based on historical data, Swansea Bay is such a site. An intensive seawater sampling programme was implemented at the designated sampling point: 0.5 h samples between 07:00 and 16:00 GMT on 60 days between 16/05/2011 and 28/09/2011, extended to 19:00 on 24 days (total 1303 samples). All FIO analyses were made in triplicate to improve the precision of the dependent variables. A coastal meteorological station was installed 3 km away and five level/discharge stations set up in local rivers and streams to derive potential environmental predictor variables. Further data were acquired from discharge gauges at larger local rivers and sewage works, plus a local tide station.

The FIO concentrations showed pronounced, consistent, within day variation of two or three $\log_{10}$ orders, a pattern not apparent in weekly compliance data. Daily rBWD water quality classification showed IE to drive the classification outcome. Daily probability of gastrointestinal illness (*p*GI) was computed from IE values and the results showed that 10% *p*GI was associated with a daily geometric mean (GM) IE of 37 cfu/100 ml.

Predictor variable matrices were constructed allowing modelling of single sample $\log_{10}$ FIO and daily $\log_{10}$ mean FIO concentrations using stepwise multiple regression. Statistically significant models for single sample results tended to provide relatively low levels of explanation (explained variance (r^2): 33% to 65%). The daily $\log_{10}$ mean FIO models produced higher levels (r^2: 55% to 89%). For both FIOs, the highest level predictors related to: solar radiation, local stream discharge and tidal variables. Turbidity in samples was also relatively important in models including this parameter. The main predictors showed plausibility in terms of slope directions.

Daily mean $\log_{10}$ FIO models were refined to predict for a 9 h time window, applicable to 4.5 h in the immediate past and future. Selected models exhibited relatively high levels of explanation (r^2: *E. coli* 81%, IE 76%), low critical misclassification (*E. coli*: 1.7%, IE 6.8% - not predicting “Poor” water quality when the observed rBWD class was “Poor”) and acceptably normal residuals distributions. Solar radiation is a dominant driver and predictions show strong diurnal patterns. Analysis of FIO concentration by time of day showed significant within-day changes, with lowest concentrations in the late morning/early afternoon. This pattern also alters rBWD classification during the day and corroborates the model predictions.

The IE model was selected for signing the site because IE: (i) generally drives the observed daily rBWD compliance outcome and (ii) provides a public health outcome in terms of *p*GI. The 37 cfu/100 ml *p*GI 10% threshold is being used for signing three times per day (twice per day on weekends) in the 2013 bathing season.

Contents

List of Figures

List of Tables

1. Introduction

Compliance monitoring for the revised European bathing water Directive (rBWD) (CEU, 2006) commenced in 2012 with the first official compliance assessment against these standards based on the 2012 to 2015 bathing seasons. Environment Agency (2012) estimates suggest that 10% of designated bathing waters in England and Wales are likely to fail to comply with the rBWD standards. Should these bathing waters consistently fail to comply by 2020, then notices prohibiting their use will be applied under the Directive. This could have potentially disastrous consequences with respect to local tourist economies. In addition, approximately 50% of the UK's current 'Blue Flag' beach awards would be lost if the numerical standards in the rBWD were applied without beach management and sample discounting as outlined by WHO (2003, 2009).

The rBWD includes provisions for this type of discounting of compliance sample results where there is a prediction and communication system in place to inform the public of impending poor water quality. The application of predictive modeling and public information is would prevent the loss of 'Blue Flag' awards and maintain access to UK bathing waters. Efforts to model FIO concentrations in bathing waters statistically have been made in the UK and world-wide using compliance data sets and data describing antecedent environmental conditions, such as rainfall, river flow and other meteorological data from existing monitoring networks (e.g. Crowther *et al.*, 2001; Francy and Darner, 2006). Some of this research has resulted in successful operational prediction systems, such as those currently used in Scotland (McPhail and Stidson, 2009; Stidson *et al.*, 2012). This report summarizes a study designed to improve statistical prediction of FIO concentrations at Swansea Bay in Wales, by enhancing the data sources used to build the statistical model beyond the available compliance sample results (typically 20 per bathing season) and existing hydrometric and meteorological networks.

The Swansea Bay location is shown in Figure 1. The beach is close to the large conurbation of Swansea, with a relatively rural hinterland. The bathing water is at risk of non-compliance with the rBWD and potential de-designation, based on the 2009-2012 compliance data. There are further urban areas around the bay, including Neath and Port Talbot, with the total urban population exceeding 250,000.

2. Materials and methods

2.1 Sampling programme

Water samples were collected from the Swansea Bay designated sampling point (DSP) at half-hourly intervals between 07:00 GMT and 16:00 GMT during three days of each week (typically Monday-Wednesday) throughout the 20 week bathing season in 2011 (16/05 2011 to 28/09/2011), following a successful trial run of the sampling and analysis protocol on 10/05/2011. This gave a total of 60 sampling days, each with 19 water quality samples. Thus, nearly as many samples were collected on individual sampling days as are collected for routine compliance monitoring at

Swansea DSP in an entire bathing season (20 weekly samples). The sampling period was extended into the evening to 19:00 GMT for 24 days between 18/07 2011 and 07/09/2011, yielding 25 samples per day.

Since the location is highly tidal, with a maximum tidal range in excess of 9 m, the sampling followed a transect, shown in Figure 1. Due to the distances involved in the intertidal zone an all terrain vehicle was required for the duration of the study. All samples were collected aseptically in sterile 1 l containers (Aurora Scientific) and immediately stored in the dark within cool boxes. Batches of samples (three per day) were then transferred to a refrigerator before dispatch to the laboratory. Samplers recorded the location details (latitude/longitude) of each sample using a hand held global positioning system (Garmin H72) and the seawater temperature using a portable electronic thermistor thermometer (Hanna Instruments 93510N). Suitable spirit-in-glass thermometers (Brannan 0-50 ˚C, 0.5 ˚C division) were provided for the event of an electronic thermometer failure.

Two samples, split from a single sample, were collected on each day. These samples were collected in a clean 5 l bucket and passed through a clean stainless steel funnel into two separate 1 l sterile bottles. The bucket and funnel were cleaned with isopropanol wipes (Vernon Carus Azo wipes) immediately before and after use. Sample numbers were randomly assigned for these samples prior to commencing the sampling programme. One sample was used as an analytical quality control (AQC) positive sample and the other for additional FIO analyses as outlined below. To determine that the procedure involved in splitting the sample was aseptic, a control of autoclaved seawater, from the site, was run through the sample splitting procedure.

2.2 Laboratory analysis

All samples were analyzed for *Escherichia coli* and intestinal enterococci (IE) using standard membrane filtration techniques. *E. coli* were enumerated using membrane lactose glucoronide agar (MLGA, Oxoid/Glycosynth) as outlined in SCA (2009, 2011). Membranes were incubated for 4 h at 30˚C, followed by 14 h at 44˚C (±0.5˚C). All green colonies were counted as *E. coli*.

At the time of the study an official analytical method for *E. coli* in UK rBWD compliance samples had not been determined and a range of methods, based on chromogenic media (e.g. MLGA), were under review. One method, subsequently adopted, uses tryptone bile glucoronide agar (TBX, Merck). A randomly pre-allocated split sample from each day was analyzed using this medium, with membranes incubated for 4 h at 30˚C, followed by 14 h at 44˚C (±0.5˚C) (SCA, 2011).

Enterococci were isolated using membrane enterococcus agar (MEA, Oxoid) by incubation for 4 h at 37˚C, followed by 44 h at 44˚C (±0.5˚C) (SCA, 2012). All maroon colonies were counted as presumptive enterococci. Membranes were then transferred to kanamycin aesculin azide agar (KAAA, Oxoid) and incubated for 6 h at 44˚C (±0.5˚C). All colonies that developed black halos were counted as confirmed IE.

All microbiological analyses were undertaken in triplicate to improve measurement precision (Fleisher and McFadden, 1980) and resulting concentrations expressed as colony forming units per 100 ml (cfu/100 ml). Serial dilutions, based on the trial run, were made using sterile Ringers solution in order to capture the appropriate range of FIO concentrations. The lower limit of detection (LLD) for *E. coli* was 3 cfu/100 ml. The theoretical LLD for IE was 1 cfu/100 ml due to the two-stage incubation. Here, a single membrane with a single presumptive colony could result, which, when incubated on KAAA does not confirm as IE. All samples were analyzed within 24 hours of collection (Mean: 10.77 h, Standard Deviation (SD): 8.12 h).

Following microbiological analysis, the samples were analyzed for turbidity (Hannah Instruments LP2000, NTU) and conductivity (Mettler Toledo SevenGo, mS). Total dissolved solids (TDS, g/l) and salinity (ppt) were also recorded using the conductivity meter.

2.3 Probability of gastrointestinal illness calculations

The World Health Organization guidelines for safe recreational waters (WHO, 2003) define water quality thresholds based on calculations related to the probability of gastrointestinal illness (*p*GI). This is outlined in Kay *et al.* (2004) and uses the relationship between IE and *p*GI from randomized epidemiological studies in the UK (Kay *et al*., 1994) to calculate *p*GI associated with the statistical distribution of $\log_{10}$ IE concentrations described by mean and SD values (Wyer *et al*. 1999). These relationships were used to calculate *p*GI associated with individual IE determinations and for each sampling day, allowing water quality comparisons to be made for defined *p*GI thresholds.

2.4 Meteorological and hydrometric monitoring

A coastal meteorological station was commissioned for the project at Black Pill, approximately 2.9 km west of the DSP (Figure 1), based on a 9 m high mast (Clark Mast ST90) attached to a building by brackets. Solar radiation sensors at the station included: global radiation (GR) (Skye Instruments SKS 1110 Pyranometer), Ultra violet-A (UVA) (315 – 380 nm, Skye Instruments SKA 420) and Ultraviolet-B (UVB) (315 – 280 nm, Skye Instruments SKA 430), all measuring W/m^2. These sensors were all mounted on a suitable bracket at 3.5 m above ground level. A combined air temperature (AT, ˚C) and relative humidity (RH, %) sensor (Rotronic HygroClip2 HC2-S3) was mounted within a radiation shield (Gill multi-plate, Model 41002) at 3.4 m above the ground surface, whilst rainfall (RF, mm) was measured using a tipping bucket rain gauge (Met One Instruments 370C 20.3 cm aperture, 0.2 mm tip) mounted on the roof apex (3.5 m above ground). A Gill Instruments WindSonic, mounted at 8.4 m above ground level, measured wind speed (WS, m/s) and direction (WD, ˚ from N). An atmospheric (barometric) pressure sensor (AP, hPa) (Met One Instruments 092) was housed inside the building alongside the data logging equipment.

The station was set to log data at 1-second intervals, using Opsis EnviLogger software. This was used to derive 15-minute averages to match hydrometric data sequences. With the exception of the GR sensor, the station was fully operational from 14:00 GMT on 16/05/2011 onwards. Data from a meteorological station at Cwm Level (Figure 1) were available as a further data source for GR, rainfall, AT, RH plus WS and WD parameters. Global radiation data from this station were used until the sensor at Black Pill was installed at 11:00 GMT on 21/06/2011. Both stations are operated by the City and County of Swansea (CCS) Environment department. A 15-minute sequence of extraterrestrial radiation (ETR, W/m^2) input (i.e. solar radiation received at the top of the atmosphere) was computed based on the latitude/longitude position of the Black Pill meteorological station. Any short periods of missing data (typically < 1 h) were replaced by linear interpolation (radiation parameters) or mean values calculated using the observations immediately preceding and following the period of missing data (all other meteorological station parameters).

The existing hydrometric monitoring network, operated by Natural Resources Wales (NRW), focuses on monitoring the major river systems in the area. The three largest rivers discharging to Swansea Bay are the Afon Tawe, Afon Nedd and Afon Afan. The closest discharge gauging stations to the outlet of each of these rivers are shown in Figure 1 (stations F to I). There are two gauging stations in the Nedd catchment, station G on the Afon Dulais tributary and station H on the main river channel. The Afon Afan also has two gauges; stations I and J (Figure 1). These stations log level (m) and discharge (m^3/s) at 15-minute intervals. Tide level data (m above chart datum) were obtained for the local NRW tidal monitoring station at Mumbles (Figure 1) as a 15-minute time series.

A further five hydrometric monitoring stations were installed for the project, based in smaller stream inputs that discharge to the bay closer to the DSP (Figure 1, stations A to E). These comprise of A. Ott Orpheus Mini pressure transducer systems set to log stream levels (m) at 5-minute intervals and report 15-minute mean values. Each unit was housed in a secure steel pipe to act as a stilling well. Two units, at stations A and C were equipped with corresponding Intelligent Top Caps (ITCs), enabling remote data logging by short message service (SMS). Standard 1 m staff gauge (stage) boards (Shelley Signs) were installed at each site. Data were processed using corresponding Hydras 3 software

A programme of open channel discharge measurements (m^3/s), across a range of stream levels, was undertaken at these stations using standard procedures (Environment Agency, 2003; Herschy, 1985). Velocities, at 0.6 of the depth from the water surface, were measured using electro-magnetic velocity meters (Sensa RC2) and the average of three measurements was recorded at each point across the channel profile. The resulting discharge measurements were related to stream level readings to generate discharge rating curves for each station.

Fifteen-minute time series from discharge monitoring stations at two sewage treatment works (STWs) were also available (Figure 1, stations K and L). Station K records the final effluent discharge (l/s) from Swansea STW, whilst station L measures the inlet flow (l/s) at Afan STW. These stations are operated by Dŵr Cymru-Welsh Water. Measurement units for these stations were converted to m^3/s to match the other discharge monitoring stations.

Any short periods of missing flow data (< 4 h) were filled in using linear interpolation. For longer periods, either catchment area scaled data from a neighbouring gauge or regression models, based on nearby gauges, were employed to replace missing values.

2.5 Statistical analysis and data preparation

With the exception of circular (angular) data such as wind direction, statistical analyses were performed using the SPSS statistical software package (version 19, SPSS 2010). The parametricity of distributions was assessed using the Shapiro-Wilk (S-W) normality test and Skewness statistic. General descriptive statistics included the mean, standard deviation (SD), range and the 95% confidence interval for the mean. Student's t-test was used to compare means between two groups. The outcome of the corresponding Levene test for homogeneity of variances was used to determine the appropriate type of t-test; based on either (i) separate or (ii) pooled variance estimates. Wind direction statistics were generated using the R statistical package (R Studio version 0.97.551), which was also used to generate corresponding wind-rose diagrams.

Robust analysis of variance (ANOVA) was employed to examine the significance of differences between more than two mean values. Here, the significance of the ANOVA is judged on (i) the Levene test for homogeneity of group variances and (ii) whether the numbers of observations (n) in groups can be considered equal. Where variances can be considered homogenous and n values are equal the significance (p) of the F statistic is used. Where n values are equal but variances are not homogenous then the Brown-Forsyth statistic p value is used. Finally, when n values are unequal and variances not homogeneous the significance of the Welch statistic is employed. The Levene test also drives the selection of an appropriate *post-hoc* test to explore the significance of multiple paired comparisons between means. Where variances are homogenous the Tukey test is used, whilst the Tamhane test is employed when variances cannot be considered homogenous.

Bivariate regression, using the SPSS curve estimation procedure, was used to develop discharge rating curves for stream gauging stations A to E (Figure 1). This allowed examination of linear and power function forms typical of such situations. The resulting functions were used to translate stream level to discharge values. Bivariate regression was also used to examine relationships between results from different FIO methods (e.g. *E. coli* by MLGA vs. *E. coli* by TBX) or, in the case of *p*GI calculation, different calculation methods. In this case, the difference of the resulting regression slope value from 1 was evaluated using Student's t-test, as outlined by Zar (2010). Microsoft Excel was used to fit non-linear curves, such as polynomial curves, with the Solver plugin applied to special non-linear cases as outlined by

Brown (2001). This plugin uses an iterative procedure to maximize the coefficient of determination (r^2) value for the specified non-linear function.

Stepwise multiple linear regression was used to explore relationships between potential environmental predictor variables (i.e. meteorological and hydrometric parameters) and FIO concentrations. For this analysis, calculations were made to generate a sequence of lagged environmental predictor variables in relation to water quality sampling times. These variables, thus, describe the antecedent environmental conditions prior to sampling.

Ten antecedent lag periods were defined for meteorological variables (0.25 h, 0.5 h, 1 h, 2 h, 3 h, 4 h, 6 h, 8 h, 10 h and 12 h before sampling). In the case of solar radiation parameters, total radiation dose received was computed for each period, with ETR, GR and UVB expressed as MJ/m^2 and UVA as kJ/m^2. A total UV dose (MJ/m^2) was computed as the sum of UVA and UVB radiation. GR was also expressed as the proportion (%) of ETR. Remaining meteorological parameters were expressed as the mean for each antecedent period with AP expressed as kPa. Calculations for mean wind direction (radians) employed circular statistical calculations as outlined by Hassan *et al.* 2009. Data for some of the antecedent periods during the first two sampling days could not be calculated for the UV parameters and AP because the relevant sensors were not active until 14:00GMT on the first sampling day (16/05/2011). This also applies to the 60-day matrix described below. Tidal variables included the tide height (m) at the time of sampling, times (h) from the temporally closest high water (HW) or low water (LW) (signed + to denote time after HW or LW and – to denote time before HW or LW) and absolute times from HW and LW. Sixteen lag periods were defined for the hydrometric rainfall and discharge variables, ranging from 0.25 h to 36 h before sampling. Intervals were as per the meteorological variables plus: 15 h, 18 h, 21 h, 24 h, 30 h, and 36 h intervals. In each case the total rainfall received (mm) and total discharge volumes (m^3) were computed for each period. Discharge variables were $\log_{10}$ transformed to reduce the absolute values involved and to reduce skew. Rainfall variables were similarly transformed, but a constant of 1 added prior transformation to account for zero values.

A similar predictor variable data matrix was developed based on data for the 60 sampling days, for which mean FIO concentrations were calculated. Variables included mean values for the seawater temperature, turbidity and salinity measured on each day. Environmental data describing the meteorological and hydrometric conditions on each sampling day were computed in a similar manner (e.g. total UVA radiation received (kJ/m^2) during the sampling day). Tidal variables included maximum and minimum tide heights and tidal range. Four antecedent lag periods were defined for this matrix (12 h, 24 h, 36 h and 48 h), each commencing at the mid-point of the sampling day and applied to all meteorological, hydrometric and tidal variables.

The forward selection stepwise regression procedure was employed to generate statistical models predicting FIO concentration (*Y*) from the environmental predictor variable matrix (X_1 to X_n). The models have the form:

$$Y = a + b_1 \times X_1 + b_2 \times X_2 + \cdots + b_n \times X_n \pm u \qquad [1]$$

where:

a = constant (intercept)
b_1 to b_n = slope coefficient for each predictor variable, X_1 to X_n
u = standard error of the estimate (stochastic disturbance term)

Two criteria for variable selection/removal were applied, the significance level (*p*-in) for a variable to enter the model was set to 0.05 and the *p*-out level, for variable removal from the model, set to 0.06 (the procedure requires *p*-out > *p*-in to operate). These settings give a 94-95% confidence window for variables in the model. In some cases, slight adjustments to this window were necessary to allow generation of a model sequence. A third criterion, tolerance, was applied to control for multicollinearity between predictor variables in a model. Tolerance is the inverse of the variance inflation factor, a low value (0.0001) allowing multicollinear variables into an equation and a high value (0.9) only allowing un-related predictors into an equation. Thus, a low tolerance model will typically contain a larger number of predictor variables, the number of which declines as tolerance increases. Model sequences were generated using tolerance values of 0.0001 plus successive 0.1 intervals between 0.1 and 0.9 and, in some cases, 0.95. This allowed analysis of consistency of variables between models and the persistence of variables as multicollinearity is increasingly controlled. The total number of predictors allowed in the models was restricted to 20.

The models were assessed using: (i) the coefficient of determination (r^2, adjusted for degrees of freedom) and (ii) the residuals distribution. r^2 defines the amount of variance in *Y* (i.e. FIO concentration) explained by the predictor variables, X_1 to X_n (i.e. antecedent environmental descriptors), in the model and indicates the overall strength of the relationship. Ideally, the distribution of residuals should be normal; this was assessed by inspection of corresponding normal-probability plots and histograms.

The overall statistical significance of all tests was evaluated at the 95% confidence level.

3. Results and discussion

3.1 Designated sampling point monitoring

A total of 1303 samples were collected and analyzed from the 60 sampling days. Two results, one for each FIO parameter, were not reported due to analytical errors. No *E. coli* were recovered from 48 samples (3.7%) and no IE from 116 samples (8.9%). Detection limit values were assigned to these samples for the purpose of statistical analysis. Descriptive statistics and normality tests showed that the FIO concentration distributions were positively skewed (skewness > 6) and demonstrated statistically significant departures from normality (S-W $p < 0.05$). Log_{10} transformation reduced skewness appreciably (< 0.2), though the distributions still showed statistically significant departure from normality (S-W $p < 0.05$). Given the reduction in skewness, the FIO data were log_{10} transformed prior to further statistical analysis.

The FIO results for individual samples are shown in Figure 2. *E. coli* concentrations ranged from <3 cfu/100 ml to 3100 cfu/100 ml (geometric mean (GM) 51 cfu/100 ml), whilst IE concentrations ranged from < 2 cfu/100 ml to 4300 cfu/100 ml (GM 31 cfu/100 ml). The striking feature of this data set is the variation in FIO concentrations within individual days, which often demonstrates a range of 2 to 3 log_{10} orders; a pattern that was continually repeated throughout the study period. The discharge record from station C (Clyne River) is shown for comparison. There appears to be a general pattern of increased FIO concentrations, and thus a decline in water quality, following hydrograph event conditions, although the daily variance is independent of antecedent rainfall.

Of the 60 samples analyzed for *E. coli* using both MLGA and TBX media, only two analyses, both using TBX, produced no *E. coli* colonies. Thus, 58 pairs were available for comparison. Figure 3 summarizes the comparison regression analysis. Student's t-test indicated that the slope of this regression (0.9476) showed no statistically significant difference from 1. This suggests that the results using both methods were comparable and that the use of MLGA in the current project would be unlikely to produce radically different results for *E. coli* than the analytical method subsequently adopted by the UK authorities (i.e. TBX).

Seawater temperatures ranged from 10.9 to 26.0 ˚C (Mean 17.2˚C, SD 2.2˚C) and were not normally distributed (S-W $p < 0.05$) though skewness (< 0.6) was low. Turbidity (Figure 4-A) ranged from 1 NTU to 3180 NTU and showed reduced skewness (< 0.4) when log_{10} transformed, though statistically significant departure from normality remained (S-W $p < 0.05$). Turbidity data were, thus, log_{10} transformed for further statistical analyses. The GM turbidity was 92 NTU. Turbidity displayed a 1 to 2 log_{10} order variation within sampling days and tended, like FIO concentrations, to be elevated following hydrograph event conditions. Distributions of all conductivity parameters showed statistically significant departure from normality (S-W $p < 0.05$) and negatively skewed distributions (skewness -2.3 to -1.9). Salinity (Figure 4-B) ranged from 23.7 ppt to 35.0 ppt (mean 32.0 ppt, SD 1.3 ppt). Lower salinity values tended to occur following hydrograph event conditions, when greater volumes of freshwater would have been present from river and stream inputs.

3.1.1 Variations in daily water quality and probability of gastrointestinal illness

The mean and SD for $\log_{10}$ FIO concentrations were calculated for each sampling day. The results are shown as GM values in Figure 5, which also shows the corresponding 95% confidence intervals for each mean. The plots show that the daily mean FIO concentrations varied considerably between days, but that the 95% confidence intervals remained relatively similar.

The daily mean and SD of $\log_{10}$ FIO concentrations were used to classify each day in terms of the rBWD standards (Table 1). The results show that daily classification is largely driven by IE concentrations, with *E. coli* driving only 3 of the 23 "Poor" overall outcomes. The results also demonstrate an effective polarization of daily outcomes at Swansea Bay, with virtually equal numbers of days in the "Excellent" (42%) and "Poor" (38%) categories. This is despite the conclusion, from compliance data, that Swansea Bay is at risk of failing to comply with the rBWD standards, with associated provisions for prohibition of bathing activities. It is, thus, critical to understand the factors affecting these observed daily rBWD outcomes.

Daily mean and SD of $\log_{10}$ intestinal concentrations were also used to calculate daily *p*GI values. The richness of the data set also allowed calculation of mean *p*GI values associated with each individual result on each day. The daily *p*GI values were then compared with the thresholds defined in the WHO guidelines for recreational waters, namely 1%, 5% and 10% *p*GI. Figure 6-A illustrates the results. Daily *p*GI values were variable, with 45% of days exceeding the upper *p*GI 0.1 (i.e. 10%) threshold, with a corresponding high risk of water associated GI (Table 2). Figure 6-B compares the *p*GI calculation methods, which show excellent parity. Student's t-test comparing the regression slope to 1 showed no significant difference from this test value. This lends an additional degree of credence to the calculations based on the daily mean and SD of $\log_{10}$ IE values.

Figure 6-C shows the relationship between the mean $\log_{10}$ IE concentration on each day and the calculated daily *p*GI. The Solver plugin for Microsoft Excel was used to fit an asymptotic logistic function, or Richard's curve, shown in the plot. This function has the form:

$$Y = A + ((K - A)/1 + e^{A-((X-M)/B)}) \qquad [2]$$

where:
X = daily mean $\log_{10}$ IE,
Y = daily *p*GI,
A is the lower asymptote (Y = 0),
K is the upper asymptote (Y = 0.3855)
and B and M are coefficients estimated in the procedure.

The value of K derives from the *p*GI calculation limits, which use a relationship between $\log_{10}$ IE and *p*GI restricted to the maximum observed IE concentration in the UK randomized controlled trials (158 cfu/100 ml) as outlined in Kay et al. (1994) and Wyer et al. (1999). The r^2 for this best fit curve was effectively 1.

Based on the corresponding daily GM IE concentrations for the WHO guideline thresholds at the Swansea Bay DSP are:

*p*GI = 0.01 : daily GM IE = 6 cfu/100 ml
*p*GI = 0.05 : daily GM IE = 20 cfu/100 ml
*p*GI = 0.10 : daily GM IE = 37 cfu/100 ml (Figure 6-C)

The daily results were split into two groups based on *p*GI, days with values > 0.1 (n = 33) and days with *p*GI ≤ 0.1 (n = 27). Using Student's t-tests, the two groups were tested for significant differences between means of daily mean $\log_{10}$ FIO concentrations (expressed as GM values) and means of daily SDs of $\log_{10}$ FIO concentrations. The results are shown in the box and whisker plots in Figure 7. Unsurprisingly, the daily mean FIO concentration associated with days with *p*GI > 10% showed statistically significant elevation compared to the *p*GI ≤10% group (Student's t-tests $p < 0.05$). This is also shown by the discrete 95% confidence intervals in Figure 7-A, which do not overlap for the individual FIOs. In contrast, Figure 7-B shows no statistically significant differences for the mean SD of $\log_{10}$ FIO concentrations in each comparison (Student's t-tests $p > 0.05$). This demonstrates that daily variance in $\log_{10}$ FIO concentrations is similar in the groups, despite the significant difference in GM FIO concentrations between the two types of days. Thus, daily variance in $\log_{10}$ FIO concentrations at the Swansea Bay DSP can effectively be regarded as constant (average daily SDs for all 60 sampling days: $\log_{10}$ *E. coli* = 0.3707, $\log_{10}$ IE = 0.4044). This finding is potentially useful in the classification of water quality using rBWD or WHO guideline criteria, which both require the mean and SD of $\log_{10}$ FIO concentrations in the relevant calculations. For example, the constant SD value for the bathing water could be applied alongside a mean value predicted using a model to determine the water quality classification for signing at the beach.

3.2 Environmental monitoring

3.2.1 Hydrometric monitoring

The level recording instrumentation at stations A to E provided 100% data capture through the study period (n = 13248). The results of discharge gauging at stations A to E are summarized in Figure 8. At least 11 discharge measurements were made at each site. Power functions were fitted to the data from four sites (r^2 adj. 0.87 to 0.98):

$$Y = b \times X^a \qquad [3]$$

where:

X = stream level (m)

Y = discharge (m^3/s)
and a and b are constants

A composite rating was fitted for station D, with a power function (r^2 adj. 0.98) fitted at stream levels below 0.36 m and a second order polynomial when level ≥ 0.36 m:

$$Y = a \times X^2 + b \times X + c \quad [4]$$

where:

X = stream level (m)
Y = discharge (m^3/s)
and a, b *and* c are constants

This function was derived from analysis of relationships between: (i) stream level (m) and channel cross sectional area (m^2) and (ii) stream level (m) and average velocity (m/s). This produced a more realistic function in terms of the physical attributes of the station site. In addition, maximum level was limited to 0.81 m, which was the level above which flow would be constrained by a bridge structure at this station. Proportions of the time series exceeding the maximum stream gauging levels ranged from < 1% at station E to 7% at station A, and was < 2% at remaining stations.

The time series for the smaller streams (stations A, B, D and E) are shown in Figure 9-A. These streams were flashy, with short hydrograph responses typical of urban streams. The corresponding record for station C is shown in Figure 2 and demonstrates a less rapid response in this larger river, the contributing catchment having comparatively large proportions of agricultural land in the headwaters (Figure 1). Table 2-A gives a statistical summary of the 15-minute discharge data through the study period. The data were highly skewed (skewness > 10).

The discharge sequences for stations F to I were complete except for 96 sequential observations at station G (09:15 GMT 25/09/2011 to 09:00 GMT 26/09/2011). The missing observations were replaced with modelled values using a lagged (+2 h) second order polynomial regression model (equation 4) predicting the discharge (Y) at station G from that at station F (X) (r^2 = 0.928). Figure 9-B illustrates the discharge at these stations and statistical summaries are given in Table 3-A. Discharges at these stations were relatively large, with maximum values at stations F and H exceeding 100 m^3/s. Hydrograph response was less rapid than observed in the smaller urban streams. The values were again highly skewed (skewness > 4.5).

Discharge at the two STWs showed a similar range (station K: 0-1.81 m^3/s, station L: 0-1.29 m^3/s) (Table 2-A). In contrast with the rivers and streams, the data for both sites exhibited low skewness (< 1). The mean discharge at station L (0.58 m^3/s) was higher than at station K (0.44 m^3/s).

The rainfall record for Black Pill was complete apart from two observations, when the station was having the global radiation sensor installed (10:30-10:45 GMT 21/06/2011). The corresponding gauge at Cwm Level indicated no rainfall at this time, so zero values were substituted in each case. The 15-minute rainfall totals recorded at Black Pill ranged from 0 to 2.2 mm (Table 3-A) and were highly skewed (skewness = 7.99).

3.2.2 Tidal and meteorological monitoring

Table 3-B illustrates the wide tidal range experienced in Swansea Bay during the study period (range: 10.23 m). The tidal data, which were complete, exhibited extremely low skew (skewness < 0.01), with a mean tide height of 5.23 m.

The mean calculated ETR during the study period was 413.35 W/m^2, with a maximum of 1166.15 W/m^2. These data exhibited low skew (skewness < 1). Global radiation data from the Cwm Level meteorological station (Figure 1) were used until the GR sensor was operational at Black Pill (n=3501). The nighttime GR data from Cwm Level were adjusted slightly, such that GR values were zero at the same time as corresponding ETR values. The data from the Black Pill GR sensor (from 11:15 GMT 21/06/2011 onwards, n=9747) were also adjusted using a baseline correction of -0.15 W/m^2, to yield zero GR during nighttime. The combined record from the two stations gave a complete 15-minute GR sequence, summarized in Table 3-B. The data were moderately skewed (skewness < 1.4). Figure 5-B illustrates the daily GR input during the study period.

The UVA and UVB sensors at Black Pill were operational from 14:00 GMT on 16/05/2011. Since no corresponding data were available from the Cwm Level station, these parameters have 56 missing records at the start of the sequence (n = 13192) amounting to < 0.5% of the record. Data from both UV sensors required consistent baseline adjustment, by -0.02 W/m^2 in the case of UVA and +0.01 W/m^2 for UVB, to yield zero nighttime values. Statistical summaries for these parameters are given in Table 3-B, and show that the energy received in the UVA spectrum is an order of magnitude higher than that in the UVB range (maxima: UVA 34.37 W/m^2, UVB 2.38 W/m^2), which are, in turn, more than an order of magnitude less than the GR input. Like GR, the UV radiation parameters exhibited moderate skew (skewness < 1.5)

Table 3-B also summarizes the other meteorological parameters measured. With the exception of AP, the sensor for which was installed along with the UV sensors, all remaining time series were complete. Air temperatures showed a range typical of summer conditions, from 5.01 to 25.93 ˚C, whilst mean RH was 80.54% (34.5% to 99.37%). Values for AP ranged between 991.2 hPa to 1034.68 hPa, typical of the fluctuating conditions between low-pressure cyclonic depressions and summer anticyclones. Wind speeds ranged between 0.22 and 12.49 m/s (mean: 2.89 m/s), with a circular mean direction of 254.15˚ (i.e. west-southwest). A further analysis of WD data is presented in Figure 10, which illustrates the predominance of winds in the westerly sector, particularly between 210˚ and 325˚. The distributions of

AT, RH and AP exhibited low negative skew (skewness < 0.7), whilst WS values demonstrated moderate positive skew (skewness < 1.5).

3.3 Multiple regression models

3.3.1 Models predicting single sample faecal indicator organism concentration

Two sets of regression models were generated for each FIO. Version 1 included measurements made at the time of sampling (position and sea temperature) and other parameters measured in the samples (turbidity and salinity) in the predictor variable matrix. These variables were excluded from the predictor matrix in the second version, since no continuous recording of these parameters could be made for a practical application of a model including them.

Table 4 summarizes the models predicting $\log_{10}$ *E. coli* concentrations based on the first version of the predictor matrix. Statistically significant models ($p < 0.05$) were generated at each tolerance level (0.0001 to 0.95). Residuals in all models all showed either slightly skewed or effectively normal distributions. Table 4 shows that the number of variables in the model progressively decreases as tolerance is increased reducing the allowed multicollinearity between predictors. The amount of explained variance, indicated by the r^2 value, progressively decreases as the tolerance increases, commencing at 0.610 (61.0%) in model 1 and ending at 0.327 (32.7%) in the final model. All eleven models included two consistent variables, entered at the first two steps. These were: (i) $\log_{10}$ discharge at station C in the 18 hours preceding a sample and (ii) global radiation received in the 4 hours prior to sampling. The first variable has a positive slope; the predicted $\log_{10}$ *E. coli* concentration increasing as discharge at station C increases. The second variable shows an inverse relationship, with the $\log_{10}$ *E. coli* concentration declining as the global radiation received increases. Both of these relationships appear plausible since rivers and streams generally exhibit elevated FIO loadings as discharge increases and increased solar radiation input has a bactericidal effect. The consistent inclusion of these variables right up to a tolerance value of 0.95 indicated that these variables were statistically unrelated to each other. Other variables that appear in a majority of the models included turbidity (positive slope, models 1-9) and salinity (negative slope, models 1-10), Tide height (positive slope, models 1-3, 5-8). These variables also appear to have plausible slope directions. For example, increased turbidity could represent increased suspended solids, which could act to shield FIOs from solar radiation and potentially act as adsorption sites for microbes. The overall levels of variance in $\log_{10}$ *E. coli* concentrations explained by the predictor variable sets were relatively low ranging from approximately one third in model 11 (4 variables) to two thirds in model 2 (20 variables). A model including the two consistent variables (i.e. steps 1 and 2) had an r^2 value of 0.300, indicating that the remaining variables entered in the models explain a relatively small amount of additional variance in $\log_{10}$ *E. coli* concentration. For instance, the 18 additional predictor variables in model 1 together account for only a similar level of explained variance (0.310) to the initial two predictors.

Table 5 shows the regression modeling results based on the second version of the predictor matrix. Overall, the levels of explained variance were reduced by excluding variables such as turbidity and salinity (r^2 range: 0.327 to 0.531). The consistent variables entered at the first two steps using the version 1 matrix remained the same in these models, model 11 in Table 5 being identical to that in Table 4.

Eleven models predicting $\log_{10}$ IE concentration in individual sample are summarized in Table 6. All were statistically significant ($p < 0.05$) and exhibited normal or only slightly skewed residuals distributions. As with the corresponding *E. coli* models, two predictors were entered consistently at the first two steps. These were: (i) GR received in the three hours to sample collection (negative slope) and (ii) $\log_{10}$ discharge at station B in the preceding 24 h (positive slope). A model including only these two predictors had an r^2 of 0.332 (Table 7, model 11) and inclusion of both variables in all models indicated that they were statistically unrelated to each other. The additional variables in the models explain comparatively low levels of additional variance. For instance, the addition of salinity to model 11 only adds a further 1.2% to the total explained variance. As with *E. coli*, variables appearing in the majority of models included turbidity (positive slope, models 1 to 9), salinity (negative slope, models 1-2, 4, 6-11) and tide height (positive slope, models 1-8, 10). Again, the relationships appear plausible. For example, the inverse relationship with salinity indicates that $\log_{10}$ IE concentration increases with reduced salinity, which might indicate elevated proportions of freshwater associated with elevated discharge, and FIO flux, from freshwater inputs during hydrograph events. Levels of explained variance (r^2 range: 0.344 to 0.603) were similar to those observed in the corresponding $\log_{10}$ *E. coli* models.

Excluding parameters such as salinity and turbidity reduced levels of explained variance (r^2 range: 0.332 to 0.536) (Table 7). The consistent variables entered at the first two steps remained, with tide height entering at the third step in all models except model 11.

Overall, the regression analysis to predict $\log_{10}$ FIO concentrations in individual samples showed:

(i) predictor variables relating to solar radiation and local stream flow contribute a high proportion of the variance explained, amounting to approximately one third of the variance in $\log_{10}$ FIO concentrations,

(ii) levels of explained variance are low (typically < 0.4), particularly for models with manageable numbers of predictors (≤ 7) for deployment in a practical context.

It is evident that explanation of variance in this FIO data set on an individual sample basis is limited, despite the enhanced precision in FIO enumeration derived from triplicate analyses and the array of data describing antecedent environmental conditions available for modelling. The resulting models, whilst interesting in terms of potential factors affecting FIO concentrations at Swansea DSP, therefore, have very limited practical utility as regulatory and/or public health protection tools.

3.3.2 Models predicting daily mean faecal indicator organism concentration

Table 8 summarizes eight statistically significant (p < 0.05) stepwise multiple regression models predicting daily mean $\log_{10}$ *E. coli* concentrations at Swansea DSP with mean sea temperature, $\log_{10}$ turbidity and salinity variables included in the predictor matrix. The models produced comparatively high levels of explained variance (r^2 range: 0.555 to 0.887), with all models explaining > 50% of the variance in the daily mean $\log_{10}$ *E. coli* concentration. However, the models exhibited either slightly skewed or skewed residuals distributions. All models included UVB radiation on the sampling day (negative slope) as the predictor entered at the first step, with seven models including the mean $\log_{10}$ turbidity (positive slope) at the second step. As a single variable, the UVB radiation received on the sampling day explained 41.4% of the variance in daily mean $\log_{10}$ *E. coli* concentration. Variables relating to local stream gauges, maximum $\log_{10}$ discharge at station E (past 36 h in models 1 -6, past 24 h model 7) and station C (past 48 h, model 8), appeared in all models, whilst variables relating to maximum tide height (on the sampling day in models 1 to 6, past 24 h in model 7) appeared in the first seven models. As with the single sample models, the slope directions for the predictors appear plausible.

Removing the predictors relating to sea temperature, turbidity and salinity resulted in five models (Table 9). Levels of explained variance were comparatively high (r^2 range: 0.624 to 0.787), with models 3 to 5 exhibiting normally distributed residuals. All models were statistically significant (p < 0.05). The UVB radiation received in the past 48 h (negative slope) was included at the first step in each model, this individual predictor explaining 42.4% of the variance in the daily mean $\log_{10}$ *E. coli* concentration. However, this variable was removed at step 4 in model 1. Maximum tide height (positive slope) on the sampling day was also included in all five models, whilst mean relative humidity (positive slope) featured in models 1 to 4, entered at step 2. With the exception of model 4, the models also included one variable relating to local stream discharge (positive slope), though the specific station was not consistent.

A summary of statistically significant models (p < 0.05) predicting daily mean $\log_{10}$ IE concentrations is given in Table 10. Seven models resulted from this analysis, with UVB on the sampling day (negative slope) entered at step 1. Explained variance exceeded 50% in all cases (r^2 range: 0.573 to 0.824) and residuals distributions were slightly skewed or normal. The individual UVB variable explained 44.0% of the variance in daily mean $\log_{10}$ IE concentration. In all models, the predictor variable entered at step 2 related to local stream discharge (positive slope). For models 1 to 6 this variable was the maximum $\log_{10}$ discharge at station E in the past 24 h (positive slope), whilst in the final model it was the maximum $\log_{10}$ discharge at station D in the past 48 h (positive slope). A tide related variable was entered at step 3 in all models, this being the maximum tide height on the sampling day (positive slope) in models 1 to 6 and the tidal range on the sampling day (positive slope) in the final model. Unlike the *E. coli* models (Table 8), turbidity was a low ranking variable only entering the first two models at step 9.

Regression models using the version 2 matrix to predict the daily mean $\log_{10}$ enterococci concentration at Swansea Bay DSP are summarized in Table 11. Levels of explained variance in these seven statistically significant models ($p < 0.05$) were, again, comparatively high, consistently exceeding 50% (r^2 range: 0.589 to 0.801). All models had slightly skewed or normally distributed residuals, except model 2. As with the models described above, UVB radiation input during the sampling day (negative slope) was entered at the first step of all models. The second step was also consistent in all models, with the maximum $\log_{10}$ discharge at station E in the past 48 hours (positive slope) entered in each model. This suggests that these two variables are statistically unrelated to each other and that this second variable contributes an additional 14.9% of the total explained variance. The maximum tide height during the sampling day was entered at the third step in models 1 to 6, adding 7.3% explained variance. The fact that the UVB, station E and maximum tide variables remain in each model up to a tolerance of 0.8 suggests that these variables are statistically unrelated to each other (i.e. exhibit low multicollinearity).

The modeling based on predicting daily mean $\log_{10}$ FIO concentrations suggests:

(i) comparatively high levels of explained variance, consistently exceeding 50% and, in some cases, exceeding 80%,

(ii) solar radiation, particularly UVB, and local stream discharge are the most important variables, with tidal variables providing some additional explanation,

(iii) the slope directions for the main predictors appear plausible,

(iii) predictor variables contributing the most to the explained variance tend to be statistically unrelated to each other.

This type of model, with associated high levels of explained variance, offers greater scope for practical deployment as part of a public information system. For example, a predicted daily mean $\log_{10}$ concentration when combined with the consistent daily SD of $\log_{10}$ FIO concentrations (described in Section 3.1.1) can be used to generate water quality classification based on the rBWD or WHO guideline criteria (Wyer *et al*., 1999).

3.3.3 Model refinement for practical application

A problem with the models described in Section 3.3.2 is that the predictor matrix includes variables matched to the sampling day (either 07:00 GMT to 16:00 or 07:00 to 19:00) and these variables appear in all models. Thus, these models would predict a mean $\log_{10}$ FIO concentration for the immediate past (i.e. 9 or 11 h). A further set of models were developed using: (i) consistent FIO data for a 9 h time window on each sampling day (07:00 to 16:00 GMT) and (ii) a revised set of sampling day variables covering a 5.5 hour time period up to the mid point of the sampling day (11:30 GMT). The sampling day tide variable was retained in the matrix, because tide for the Mumbles station can be calculated reliably. All other lag periods were also fixed to the 11:30 GMT mid-point in the predictor matrix, to reflect the consistent sampling day. These models would, thus, predict a mean $\log_{10}$ FIO concentration for the mid point of a 9 h time window, valid for 4.5 hours in the immediate past and future.

A further test for these models was set in terms of corresponding rBWD water quality misclassifications. The critical misclassification (C. M.) in this context was judged to be cases where the model predicted water quality that was not "Poor" (i.e. it was "Excellent", "Good" or "Sufficient") when the observed classification was actually "Poor". This situation would be regarded as not protective of public health, since a corresponding sign would indicate acceptable water quality. Other outcomes, for example the model predicting "Poor" water quality when the observed classification was "Good", were considered precautionary (i.e. the "Poor" water quality signing would be protective of public health even though the predicted outcome was incorrect). For this analysis the observed and predicted rBWD outcomes were assessed using the average SD of $\log_{10}$ FIO concentrations for all 60 9 h periods (*E. coli*: 0.3657, IE 0.3986) to compute the relevant geometric 90%ile and 95%ile values to compare with the rBWD criteria. The proportion (%) of C. M.s was then computed for each model and are shown in Tables 12 and 13.

Models predicting daily mean $\log_{10}$ *E. coli* concentration for the 9 h time window are summarized in Table 12. All models were statistically significant ($p < 0.05$) and explained over 50% of the variance in 9 h mean $\log_{10}$ *E. coli* concentration (r^2 range: 0.566 to 0.802). Residuals distributions were variable, with models 3 and 6 exhibiting particularly skewed distributions. As with previous models, variables relating to solar radiation (UV, negative slope) and local stream discharge (station E, positive slope) were entered at the first two steps. These variables were subsequently removed at later steps in some models, as indicated in Table 12 (models 1 – 4), along with an RH variable. Replacement solar radiation (GR, negative slope) and local stream discharge (station C, positive slope) variables were included in models 2 – 4, retaining the solar radiation and local stream elements in these models. All models included a tidal variable (positive slope).

Generation of a model sequence predicting 9 h mean $\log_{10}$ IE required adjustment of the regression criteria. The p-in value was raised from 0.05 to 0.051, giving a confidence window for variables in the model of 94% to 94.9%. Without this change, the same model resulted for all tolerance values between 0.1 and 0.8. This slight adjustment resulted in eight distinct models, summarized in Table 13. All were statistically significant ($p < 0.05$) and explained more than 50% of the variance in the dependent variable (r^2 range: 0.541 to 0.765). The first two steps in all models involved local stream discharge (station E, positive slope) and solar radiation related variables (UVA, models 1 – 7, GR, model 8, both negative slope), though the local stream variable was removed at a later step in some models (models 2 and 3), following inclusion of a different local discharge variable associated with station C (positive slope). Models 1 to 7 consistently included a tidal variable (positive slope).

A "best" model for each FIO was selected from Tables 12 and 13. For *E. coli*, model 2 was selected as it exhibited: (i) high explained variance (81.4%), low rBWD C. M. (1.69%) and (iii) a relatively normal residuals distribution (Figure 11-C). A plot of observed and predicted values using this model is shown in Figure 11-B, whilst Figure 11-A displays a temporal plot of the running 9 h GM *E. coli* concentration predicted by the model during summer 2011. This plot also shows the correspondence of the prediction with the observed 9 h GM concentrations from the sample days.

The corresponding IE model selected was model 3 from Table 13, which displayed: (i) the lowest rBWD C.M. (6.78%), (ii) relatively high explained variance (75.9%) and (iii) normally distributed residuals (Figure 12-C). Figure 12-A illustrates the behaviour of this model as a time series of 9 h running GMs through the summer of 2011 along with the observed sample day GMs.

The FIO concentration data from the 24 days on which sampling was extended to 19:00 GMT were used to calculate six successive 9 h mean $\log_{10}$ values for each 0.5 h interval between 07:30 – 16:30 GMT and 10:00 – 19:00 GMT. These 144 observed 9 h mean $\log_{10}$ FIO values were compared with the corresponding model predictions, as a form of validation. The results are shown in Figure 13, which shows distinct groups of points associated with individual days. In some cases these show positive trends, with observed and predicted values increasing. Other cases are static, with the observed values increasing slightly compared to the predicted values (vertical lines of points), whilst in other cases the predicted values decline whilst observed values increase (negative slope). Both models tend to under-predict at the top end of the observed range of values. The IE model also allows assessment of *p*GI outcome, based on the *p*GI 0.1 threshold IE concentration (37 cfu/100 ml) shown in Figure 13-B. Based on the 144 observations shown, 124 (86.11%) showed the correct classification with respect to the threshold. However, 12.5 % of cases have observed values > 37 cfu/100 ml with corresponding predictions below this threshold (shown as red points in Figure 13-B). This represents a critical misclassification, where the model would not result in a "Poor" water quality sign when the observed water quality indicates otherwise. A further 2 cases, just 1.39%, showed "Good" observed water quality but "Poor" predicted water quality. These are regarded as protective of public health.

3.3.4 Within-day variation of faecal indicator organism concentrations

It is evident from Figure 2 that there was considerable within day variation in FIO concentrations at the Swansea bay DSP. The multiple regression modelling has suggested that solar radiation related variables are important predictors of FIO concentration and the model predictions, illustrated in Figures 11 and 12, suggest a strong diurnal pattern in the running GM FIO concentration, driven by the solar radiation input predictor variables. With this in mind, a temporal analysis of the FIO concentration data was undertaken.

The 07:00 -16:00 GMT data from all 60 days were grouped into two periods: (i) samples collected between 07:00 GMT and 11:00 and (ii) samples collected between 11:30 and 16:00 GMT. The days were further classified according to the calculated *p*GI, based on the same 07:00 – 16:00 GMT data, split into days with *p*GI ≤ 0.1 (34 days) and *p*GI > 0.1 (26 days). This split is slightly different to that based on all available data for each day (i.e. including 16:30 – 19:00 GMT data) shown in Table 2 (*p*GI ≤ 0.1: 33 days, *p*GI > 0.1: 27 days). The results are shown in Figure 14. Using data for all 60 days, the GM FIO concentrations for the late morning-afternoon period were significantly lower than the earlier morning period (Student's t-test $p < 0.05$). ANOVA of the four groups, based on time and *p*GI, revealed statistically significant differences between group GMs for all comparisons (Tamhane multiple comparison test $p < 0.05$). This demonstrates that pattern of late morning-afternoon reduction in GM FIO concentrations, compared to the earlier morning, persists even when water quality has significantly deteriorated (i.e. on days with *p*GI > 0.1).

The data from the 24 days with 07:00 – 19:00 GMT samples were split into three temporal categories: (i) 07:00 and 11:00 GMT samples, (ii) 11:30 – 15:00 GMT samples and (iii) 15:30 – 19:00 GMT. The days were also classified based on calculated daily *p*GI : (i) 11 days with *p*GI ≤ 0.1 and (ii) 13 days with *p*GI > 0.1. The results are shown in Figure 15. Based on all 24 days, the GM *E. coli* concentrations in the late morning-early afternoon period was significantly lower than during the earlier morning (Tamhane multiple comparison test $p < 0.05$, Figure 15-A). No other comparisons were significantly different for *E. coli*. This suggests that, on average, the GM *E. coli* concentration in the late afternoon-early evening recovered to a similar level as the earlier morning samples. The pattern for IE (Figure 15-B) was more distinct, with the late morning-early afternoon GM being significantly lower than both other groups, which were not significantly different from each other.

The ANOVA for the groups split by time period and *p*GI class showed GM FIO concentrations to be significantly elevated in the *p*GI > 0.1 group for all time periods. Comparisons of GM *E. coli* (Figure 15-A) showed no statistically significant differences between the three time periods on days with *p*GI ≤ 0.1 (Tamhane multiple comparison test $p > 0.05$). In contrast, the GM IE concentration in the late morning-early afternoon was significantly lower than the earlier morning in the *p*GI ≤ 0.1 group (Tukey multiple comparison test $p < 0.05$). No other comparisons were significantly different in this *p*GI group. Comparisons of GM *E. coli* concentrations in the *p*GI > 0.1 group showed the only significant difference to be between the morning and late morning-early afternoon groups (Tukey multiple comparison test $p < 0.05$, Figure 15-A). Similar comparisons between GM IE values in the *p*GI > 0.1 group showed the late morning GM to be significantly lower than both the earlier morning and late afternoon-evening groups, which were not significantly different from each other (Figure 15-B). Figure 16 shows the pattern of GM FIO concentrations by time of day for all days and the two *p*GI groupings. The pattern based on all days suggests that, on average, FIO concentrations decline through the day to early afternoon and increase during the late afternoon and early evening. The patterns for days with *p*GI > 0.1, with corresponding elevated FIO concentrations, appear more exaggerated, whilst the *p*GI ≤ 0.1 days exhibit a more subdued pattern.

These results suggest significant within day changes in FIO concentration that have a diurnal pattern, with the lowest FIO concentrations associated with the late morning-early afternoon period. On average, based on the 60 day 07:00 to 16:00 GMT data, the resulting rBWD *E. coli* classification would shift from “Sufficient” in the earlier morning to “Good” in the late morning-early afternoon, whilst the shift based on IE is from “Poor” to “Sufficient” (Table 14-A). Using the three classes applied to the 24 days with sampling extended to 19:00 GMT, the rBWD *E. coli* classification is, on average, “Good” in the late morning-early afternoon, and “Sufficient” in both the earlier morning and late afternoon-early evening periods (Table 14-B). For IE the change in GM concentrations between the three periods is enough to alter the corresponding rBWD classification from “Poor” in the earlier morning to “sufficient” in the late morning-early afternoon, with a return to the “Poor” classification in the late afternoon-early evening (Table 14-B). Table 14-C shows the diurnal change in rBWD classification based on values for each hour on an average sampling day. This observed, dynamic, within day pattern of FIO concentration change lends credence to the diurnal patterns suggested by the model predictions (Figures 11 and 12). This diurnality has clear implications with respect to compliance sampling regimes, which are often driven by factors such as sample delivery times. It also presents considerable challenges to predictive modelling efforts, which may currently produce an advisory applied to an entire day. The observations from the current study suggest that models should attempt to account for within day variations in FIO concentrations to: (i) provide timely public information on changes in water quality and (ii) sign a site appropriately for as long as required.

3.3.5 Model application

The two selected models had five common variables (and data sources):

(i) maximum tide (m) on sampling day (i.e. 9 h window),
(ii) ETR (MJ/m2) in the past 24 h,
(iii) $\log_{10}$ maximum discharge (m^3) at station C in the past 48 h,
(iv) $\log_{10}$ discharge (m^3) at station I in the past 48 h,
and
(v) mean wind speed (m/s) in the past 24 h.

For the *E. coli* model, three further meteorological variables were involved:

(i) GR (MJ/m^2) in the past 12 h
(ii) mean WD (radians) in the past 12 h
and
(iii) mean AP (kPa) in the past 12 h,

whilst the selected IE model included two further variables:

(i) UVA radiation (kJ/m^2) in the past 12 h and
(ii) $\log_{10}$ maximum discharge (m^3) at station G in the past 5.5 h

Thus, running the two models would require data input streams from 10 separate sources, eight of which would be from continuous logging sensors (gauging stations (3) and meteorological parameters (5)), whilst two (ETR and tide) could be calculated in advance.

The rBWD classifications for the 60 sampling days showed that IE dominated the overall classification outcome (Table 1), with *E. coli* driving the "Poor" water quality classification in 3 of 23 instances. Unlike the *E. coli* model, the IE model could also be related to a public health risk outcome in terms of *p*GI, as discussed in Section 3.1.1 (Figure 6). It was, thus, decided to concentrate operational efforts on a single model, predicting 9 h mean $\log_{10}$ IE concentration (Figure 12) referenced to *p*GI outcome. This reduced the required data input streams to five sensors (3 gauging stations (C, G and I) and 2 meteorological sensors (UVA and WS)) plus the two pre-calculated input parameters (ETR and tide). It was thought that this would have some practical benefits in terms of potential data loss from factors such as instrumentation and data transfer failures.

At the time of writing (summer 2013), the model has been run operationally for 11 weeks using Excel workbooks, the first of which is used to collate the input data, which are then used to populate a second workbook. This workbook transforms the data, in terms of units, and performs the relevant calculations to predict the running mean $\log_{10}$ IE value at a 0.25 h interval using the model. This value is then used to generate corresponding rBWD and WHO guideline classifications. The value is also compared to the 0.10 *p*GI threshold generated using the relationship between mean $\log_{10}$ IE and *p*GI (Figure 6-C). Exceedence of the 0.1 (i.e. 10%) *p*GI threshold (GM IE 37 cfu/100 ml – Figure 6-C) is used as the basis for manual signing at the DSP. Model results for the 2011 bathing season are shown in Figure 12-A in relation to this threshold. Three signs are used: (i) Good water quality is predicted, (ii) Poor water quality is predicted and (iii) prediction unavailable. The latter is used in the event of the model not being able to be run, due, for example, to equipment failure.

Data are received from three gauging stations three times per day (just after 09:00, 12:00 and 15:00 GMT) and relevant data from the meteorological station are retrieved from a file transfer protocol (FTP) data stream at the appropriate time (just after 09:15, 12:15 and 15:15 GMT). The model uses data from station J as surrogate for station I, due to refurbishment work at station I. Station J is 3.2 km upstream of Station I on the same river (Figure 1) and a comparison of past records shows excellent agreement between the stations. The data from the two NRW gauges (stations G and J, Figure 1) are currently received by e-mail, whilst the data from the local river (station C, Figure 1) arrive by SMS and are processed, and extracted, locally using relevant software (A. Ott Hydras 3 and Hydras 3 Rx packages). So far, data capture has been very good, with few equipment/data transfer failures impacting on the signing programme. Experiments are currently underway to transfer the station C data by FTP, and early results, using a second Orpheus Mini level logger, are encouraging.

The model is currently run three times per day during the working week and the DSP signed accordingly. The resulting sign is also displayed on the CCS website, via a "Twitter" feed (http://www.swansea.gov.uk/index.cfm?articleid=29433). The model is also run twice per day at weekends (12:00 and 15:00 GMT). Figure 17 shows two examples of operational plots showing the behaviour of the predicted gunning GM IE concentration and the predictor variables during two contrasting weeks during the 2013 bathing season. The predicted GM line is compared to the threshold GM 37 cfu/100 ml concentration used for signing; values above the line indicating a "Poor" water quality prediction and values below indicating "Good" predicted water quality. In the first week (Figure 17-A) "Poor" water quality was predicted for approximately 36 hours between 24 h and 60 h. This corresponded to rainfall driven hydrograph events (shown in the traces for the three gauging stations) and corresponding depression of UVA input, associated with cloudy skies. Towards the end of the week the river discharge pattern had settled to lower levels and maximum UVA input had increased, minimizing the time that "Poor" water quality was predicted. The maximum tide level variable also decreased through this week, which also acted to reduce the predicted GM IE levels. The plot for week 9 (Figure 17-B) corresponded with a summer anticyclone, with clear skies. This is shown in the consistent UVA pattern through the whole week. The river discharge variables showed only slight reductions through the week and maximum tide levels were also moderate through out the period. This resulted in a consistent diurnal cycle of running GM IE concentrations through the week, with values exceeding the 37 cfu/100 ml threshold for short periods during the nighttime.

There are future plans to automate the system by incorporation into a 'Nowcast' air quality prediction system linked to electronic signs. This will be used to make running hourly predictions, to account for within day changes in *p*GI. This should: (i) ensure that the DSP is signed appropriately for as long as required and (ii) inform the public of likely "Poor" water quality, with a high associated GI risk, as soon as possible. A programme of confirmation (closure) and replacement sampling has also been instigated alongside the compliance monitoring at the Swansea Bay DSP. This is based on the sign displayed at the time a compliance sample is taken. If "Poor" water quality is predicted then a confirmation sample is taken within 72 h of the compliance sampling time. A replacement compliance sample is then taken within 7 days of the confirmation sample. At Swansea Bay, this should alleviate the impending threat of prohibition should this bathing water fail to comply with rBWD up to 2020.

4. Summary and conclusions

1. An intensive programme of monitoring faecal indicator organism (FIO) concentrations at Swansea Bay designated sampling point (DSP) was successfully implemented during the summer of 2011. Seawater samples were collected at 0.5 h intervals between 07:00 and 16:00 GMT on 60 days through the 20-week bathing season (16/05/2011 and 28/09/2011). Sampling was extended to 19:00 GMT for 24 days between 18/07/2011 and 07/09/2011. A total of 1303 samples were collected.

2. Samples were analyzed for *E. coli* and intestinal enterococci (IE) (colony forming units (cfu)/100 ml) using standard membrane filtration methods and in triplicate. Only two results, one for each FIO, were missing, due to analytical errors. The resulting concentrations were found to exhibit closer approximation to normality when $\log_{10}$ transformed.

3. The FIO concentrations showed pronounced, consistent, variation within individual sampling days, often amounting to two or three $\log_{10}$ orders of magnitude. Patterns also showed elevated FIO concentrations following hydrograph event conditions in the rivers. Daily classification using the criteria of the revised European bathing waters Directive (rBWD) showed IE, rather than *E. coli*, was the principal driver of water quality classification.

4. The IE data were used to calculate the probability of gastrointestinal illness (*p*GI), as used in the derivation of WHO guideline standards for recreational waters (Kay *et al.*, 2004). These results were used to assess of modeling outcomes in terms of public health. For example, a relationship between daily mean $\log_{10}$ IE and *p*GI was developed and used to generate threshold water quality values for relevant *p*GI values (e.g. *p*GI 0.1 – used to define high GI risk – has a corresponding geometric mean (GM) IE concentration of 37 cfu/100 ml).

5. An analysis of within-day variation, measured by the daily standard deviation (SD) of $\log_{10}$ FIO concentrations, demonstrated that the mean daily SD on days with high *p*GI (> 0.1) was not significantly different from that on days with low *p*GI (≤ 0.1). Daily variation in $\log_{10}$ FIO concentrations at Swansea Bay can, thus, be regarded as effectively constant.

6. A meteorological station was installed at a coastal location approximately 3 km west of the DSP. This measured rainfall, solar radiation input (Global, UVA and UVB), air temperature, relative humidity, atmospheric pressure, wind speed and direction. Parallel data were also available from an inland station.

7. A network of five stream level recorders was installed in local rivers and streams. Discharge rating curves were developed for these sites through a programme of discharge measurements. Further data were available from the gauging station network on the larger rivers discharging to the bay and tide level data were also available from a local tide gauge.

8. A virtually complete set of 0.25 h time series were generated for the meteorological, tidal and gauging station data covering the water quality monitoring period. UV radiation data were missing at the start of the period because the sensor was not installed until the first sampling day. Other short periods of missing data were substituted by interpolation or regression modelling, based on data from neighbouring stations.

9. Matrices of antecedent environmental predictor variables were generated from the meteorological, tidal and gauging station data. Separate matrices were developed for statistical modelling of: (i) $\log_{10}$ FIO concentrations in individual samples and (ii) daily mean $\log_{10}$ FIO concentrations using stepwise multiple regression.

10. The individual sample models, though statistically significant, exhibited low levels of explained variance (typically between 33% and 65%). The most important predictors in these models were solar radiation, local river and stream flow and tidal variables. Turbidity was also important, when included in the models.

11. Models predicting daily mean $\log_{10}$ FIO concentrations were more successful, with maximum levels of explained variance approaching 90% and always exceeding 50%. Again, important predictors were related to solar radiation, local stream flow and tide. Turbidity was an important predictor in *E. coli* models when included, but not the IE models.

12. The daily mean $\log_{10}$ FIO models were refined to provide predictions for a 9 h time window, with antecedent lag periods tied to the mid-point. This produced models to predict the mean applicable to 4.5 h in the immediate past and future. These models provided between 54% and 80% explained variance, with main predictors related to solar radiation, local stream flow and tide.

13. Two models, one for each FIO, were selected based on: (i) level of explained variance, (ii) distribution of residuals (normality) and (iii) critical misclassification (predicting “Good” water quality when observed water quality was “Poor”, based on rBWD). Both models exhibited relatively high levels of explained variance (*E. coli*: 81%, IE: 76%), low critical misclassification (*E. coli*: 1.7%, IE 6.8%) and acceptable residuals distributions. Important predictors were again related to solar radiation, local stream flow and tide.

14. Predictions from both models suggested pronounced diurnality in the running GM FIO concentration sequences, related to the solar radiation predictor. Temporal analysis of the FIO data demonstrated statistically significant within day temporal variation in FIO concentrations, with minimum values in the late morning-early afternoon compared to the earlier morning and late afternoon early-evening periods. On average, this variation was enough to produce within day variation in rBWD classification and supports the model results. This within-day variation of FIO concentrations has implications with respect to compliance monitoring and provides challenges for existing FIO modelling strategies, which generally seek to predict a

“bathing-day” water quality. The empirical data acquired in this project suggest that the bathing day does not, in reality, exhibit a uniform water quality. Hence, the existing models world-wide that are based on this assumption may, therefore, provide an inappropriate and potentially dangerous simplification of reality.

15. Given that IE tends to drive the daily rBWD classification and can, unlike *E. coli*, be applied in a public health context, the IE model was adopted for practical application for signing at the Swansea Bay DSP. Signing is based on a threshold GM of 37 cfu/100 ml, above which the *p*GI exceeds 0.1. Applying a model that provides within day prediction ensures that the DSP is appropriately signed for as long as required and informs the public of “Poor” water quality as soon as possible.

16. The selected model has been applied manually during the 2013 bathing season. It is currently run three times each day during the working week and twice per day at weekends. Appropriate signs are then displayed at the DSP. A complementary confirmation/replacement compliance sampling programme has also been put in place, linking the model outcome to compliance monitoring.

17. Overall, this successful modelling exercise has demonstrated that prediction of FIO concentrations at a bathing water is possible given: (i) collection of a sufficiently rich data set describing water quality and (ii) a robust set of predictor variables. This suggests that modeling exercises cannot rely on compliance data sets and existing environmental monitoring networks, especially with the challenge to model within day variations in water quality. For example, the most important predictors in the current project derive from sensors specifically installed for this work, namely solar radiation (UVA) and a local river gauge.

Acknowledgements

This research project could not have been completed without the assistance of many colleagues at the City and County of Swansea, Natural Resources Wales, Dŵr Cymru-Welsh Water and Aberystwyth University. We are indebted to: Jenni Aazem, Elin Chapman, Richard Colman, Emyr Evans, Sîan Hulme, Keith James, Adrian Johnson, Sarah Jones, Vicky Morgan, Shayne Perry, Tom Price, Emma Smith and Tom Sohal (City and County of Swansea, Environment Department) who were involved in the seawater sampling under all weather conditions; Andrew Dempster, Duncan Ewin, Adam Fisher, Zeeshaan Jamil and Stefan Trimble (Aberystwyth University) who provided expert technical support for the microbiological analyses; Deidre Williams, Joe Pearce and Ieuan Davies (Aberystwyth University) who provided technical support with hydrological measurements; Paola Rabaiotti, Paula Hopkins, Dave Kelly, Darren Morgan and Dave Whitehead (Aberystwyth University) who provided clerical and logistical support; Jon Jones, Chris Palmer and Martin Haywood-Samuel (Natural Resources Wales) who provided support with project planning and data acquisition and Fergus O’Brien and Rob Jenkins (Dŵr Cymru-Welsh Water) for providing sewage treatment works discharge data. We also thank Robin Guy and Simon Wills (Ott Hydrometry) for their technical support and advice.

References

Brown, A. (2001) A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Computer Methods and Programs in Biomedicine* 65: 191-200.

Council of the European Communities (CEC) (1976). Council Directive of 8 December 1975 concerning the quality of bathing waters. *Official Journal of the European Communities No L 31 5.2.1976*, p1-7.

Council of the European Union (CEU) (2006). Directive 2006/7/EC of the European Parliament of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. *Official Journal of the European Union No L 64 4.3.2006*, p37-51.

Crowther, J., Kay, D and Wyer, M.D. (2001). Relationships between microbial water quality and environmental conditions in coastal recreational waters: the Fylde coast, UK. *Water Research*, 35(17): 4029-4038.

Environment Agency (EA) (2003). *Draft Hydrometric Manual. Chapter 4: Instantaneous flow measurement*. Environment Agency, Bristol.

Environment Agency (2012). *Projected revised Bathing Water Directive classifications using 2009-12 monitoring data for bathing waters in England and Wales.* 12 pp. Environment Agency.
http://webarchive.nationalarchives.gov.uk/20130221161545/http://archive.defra.gov.uk/environment/quality/water/waterquality/bathing/documents/projected-rbwd-classifications.pdf

Fleisher, J. M. and McFadden R. T. (1980). Obtaining precise estimates in coliform enumeration. *Water Research* 14: 477-83.

Francy, D. S. and Darner, R. A. (2006). *Procedures for developing models to predict exceedances of recreational water-quality standards at coastal beaches.* Techniques and methods 6-B5. 34 pp. United States Geological Survey.

Hassan, S. F., A. G. Hussin and Zubari, Y. Z. (2009). Analysis of Malaysian wind direction data using ORIANA. *Modern Applied Science* 3(3): 115-119.

Herschy, R. W., 1985. *Stream flow measurement*. 547pp. Elsevier Applied Sciences Publishers, London and New York.

Kay D., Fleisher J. M., Jones F., Salmon R. L., Wyer M. D., Godfree A. F., Zelanauch-Jacquotte Z. and Shore R. (1994). Predicting the likliehood of gastroenteritis from sea bathing: results from randomised exposure. *The Lancet* 344: 905-909.

Kay, D., Bartram, J., Prüss, A., Ashbolt, N., Wyer, M. D., Fleisher, J. M., Fewtrell L., Rogers, A. and Rees, G. (2004). Derivation of numerical values for the World Health Organization guidelines for recreational waters. *Water Research*, 38(5): 1296-1304.

McPhail, C. and Stidson R. (2009). Bathing water signage and predictive water quality models in Scotland. *Aquatic Ecosystem Health and Management*, 12(2): 183-186.

SPSS (2010). *IBM SPSS Statistics 19 Brief Guide.* 161pp. SPSS inc.

Standing Committee of analysts (SCA) (2009). *The microbiology of drinking water (2009) – Part 4 – Methods for the isolation and enumeration of coliform and Escherichia coli (including E. coli O157:H7). Methods for the examination of waters and associated materials.* Environment Agency, 101 pp.

Standing Committee of analysts (SCA) (2011 - Draft). *The microbiology of recreational and environmental waters (2011) – Part 3 – Methods for the isolation and enumeration of coliform and Escherichia coli (including E. coli O157:H7). Methods for the examination of waters and associated materials.* Environment Agency, 111 pp.

Standing Committee of analysts (SCA) (2012). *The microbiology of drinking water (2012) – Part 5 – Methods for the isolation and enumeration of enterococci. Methods for the examination of waters and associated materials.* Environment Agency, 31 pp.

Stidson, R. T., Gray, C. A. and McPhail, C. D. (2011). Development and use of modeling techniques for real-time bathing water quality predictions. *Water and Environment Journal* 26: 7-18.

World Health Organization (WHO) (2003). *Guidelines for Safe Recreational-Water Environments Volume 1: Coastal and Fresh-Waters*. 219 pp. World Health Organization, Geneva, Switzerland.

World Health Organization (WHO) (2009).. Addendum to the WHO Guidelines for safe recreational waters, Volume 1, coastal and fresh waters. List of agreed updates. 33 pp. World Health Organization, Geneva, Switzerland.

Wyer M. D., Kay D., Fleisher J. M., Salmon R. L., Jones F., Godfree A. F., Jackson G. and Rogers A. R. (1999). An experimental health-related classification for marine waters. *Water Research* 33 (3): 715-722.

Zar, J. H. (2010). *Biostatistical Analysis*. Fifth Edition. Pearson Prentice Hall. 944pp.